Proposal for

The Middlebury Initiative for Data and Digital Methods MiddData

August 2020

Submitted to the Office of the Provost by

Caitlin Knowles Myers, John G. McCullough Professor of Economics
Kathryn Morse, Professor of History, Elder Chair in Environmental Studies
Mike Roy, Dean of the Library

With contributions from

Carrie Anderson, Assistant Professor of History of Art and Architecture
Amy Collier, Associate Provost for Digital Learning
Bill Koulopoulos, Director of Technology and Learning Spaces
Sarah Laursen, Assistant Professor of History of Art and Architecture
Matthew Lawrence, Assistant Professor of Sociology
Alex Lyford, Assistant Professor Mathematics
Vijay Menta, Associate Vice President and Chief Information Officer for ITS
Jason Mittell, Professor and Chair of Film and Media Culture
Amy Morsman, Professor of History and Director of the DLA
Phil Murphy, Associate Professor at MIIS and Director of the META Lab
James Ralph, Dean of Faculty Development and Research
Daniel Scharstein, Charles A. Dana Professor of Computer Science
Wendy Shook, Science Data Librarian
Mark Spritzer, Associate Professor and Chair of Biology

Summary

We envision a future in which every Middlebury student is data and digital fluent. To accomplish this goal, we propose a major new initiative to provide equitable and inclusive access to powerful tools for empirical research, data analysis, and critical digital scholarship from the moment students arrive on campus. The **Middlebury Initiative for Data and Digital Methods (MiddData)** would prepare students to evaluate data and data-driven arguments from a variety of perspectives, discover insights from data both independently and in collaboration with others, form their own data-driven arguments, and effectively communicate their findings through a variety of media.

We imagine MiddData as a space where students, faculty, and staff from across Middlebury gather and analyze the big, messy, and imperfect "real world" of data and digital information in search of answers to the questions that fuel their curiosity. To help facilitate that vision, the College would create campus-wide introductory courses in data science, digital methods, and statistics accessible to a broad audience and geared towards incoming first-year students. Students enrolled in these introductory courses could additionally enroll in associated "connector courses" (existing and new) in which they would apply the methodological tools they acquire to inquiry specific to a discipline of their choosing. In addition, MiddData would help expand data and digitally-oriented courses across the curriculum including research-based practicums and data and digital "boot camps." MiddData also would support the creation of a new minor in Data Science and would explore the possibility of a similar minor in the Digital Liberal Arts. Two faculty co-heads, one with a background in the digital humanities and the second with a background in statistics, computer science or a related field, would lead this new initiative.

In this proposal we identify and discuss challenges and opportunities confronting Middlebury to make the argument for this new initiative. We then enumerate proposed innovations offered by the MiddData Initiative. For interested readers, an appendix details the inclusive process that has produced this proposal.

Challenges and opportunities: Why do we need a new initiative?

All of our students need to be fluent in data and digital methods, and aware of the ways in which data and digital information shape the production of knowledge. Only with such fluency can they fully participate in an increasingly digital landscape and data-centered job market, and meet Middlebury's educational missions across the broader institution. As we prepare students to be rigorous thinkers and effective communicators who solve the world's most challenging problems, we must train them to evaluate data-driven arguments, make informed choices about which tools and methods are appropriate, construct meaning and arguments from data and other digital forms, and effectively communicate their findings and insights via digital tools and methods.

Our moment calls with particular force for a concentrated commitment to innovation in the realm of data and digital knowledge. In her new podcast, *The Last Archive*, historian Jill Lepore investigates the question: "Who Killed Truth?" As citizens, students, professionals, and scholars, we are awash in data and digital information about atmospheric CO_2 levels, COVID-19 infection rates, rates of police shootings, and so many other pressing data-producing challenges. At the same time, as Lepore points out, few of us have ever felt so far away from knowledge agreed upon as true.

As educators, we face formidable challenges: how do we educate our students to understand the practice of rigorous empirical research and the production of reliable knowledge, while also recognizing the complex realities of terms like "reliable knowledge" and even "truth"?

One necessary step in meeting this challenge is to gather scholars and teachers from across the institution to collaborate. Our colleagues in math, the natural sciences, the social sciences,, and other quantitative fields are developing new methods for producing and analyzing big data at new scales, with vast new powers to discern patterns out of nearly infinite reams of information.

At the same time, scholars and teachers in the humanities and arts find their archival research transformed by vast efforts to digitize primary sources and art works around the globe. The amount of archival material available by internet search expands every day. With scholars increasingly pursuing research through access to global databases of sources, and working to create digital archives of their own, we face the opportunity and the challenge to transform our students' skills in digital methods of research and communication.

Our most challenging opportunity lies in cultivating collaborations in which teachers and scholars across the disciplines draw from each others' strengths and methods to enrich their own and each other's work. How can humanists use quantitative methods to ask and answer compelling literary, historical, and philosophical questions? How can economists and field

biologists draw quantitative information from historical records and provide new insights about the past and present for historians to then use to re-evaluate historical understanding?

Beyond such cross- and inter-disciplinary work, how can humanists and social scientists share their skills in qualitative cultural analysis to help quantitative scholars advance their use and understanding of data, its sources, and the cultural framework of its production? Digital Archivist Trevor Owens at the Library of Congress writes that "Data is always manufactured...always, at least indirectly, created by people."[1] How, on the other hand, can artists and historians enlist data scientists to infuse creative work and historical narratives with quantitative representations of change over time? Our community has an impressive track record of just such collaborations, but it is often the result of just one or two scholars connecting beyond the boundaries of institutional structures and funding. We thus find ourselves with the timely opportunity to build a new institution-wide community focused on data and the digital. However, Middlebury currently faces significant barriers to access to implementing and teaching empirical data science and digital methods.

Overwhelming demand from students and insufficient supply of faculty: While a talented group of Middlebury faculty already teach courses across the curriculum covering data collection, management, visualization, analysis, and communication, these overburdened and non-competitively paid colleagues face rising student demand, leaving little time for innovation or developing optimal pedagogical strategies. At the same time, the administration struggles to balance the distribution of faculty lines between the large departments with surging enrollments, and smaller departments that nevertheless must maintain a minimum number of faculty to achieve sufficient curricular diversity. These frictions can make it difficult to efficiently and equitably respond to rapidly changing enrollment patterns.

Disciplinary silos: Despite our different disciplinary questions and perspectives, we see a convergence among the tools and methods we use. We can teach and support these tools and methods broadly, across departments and campuses. For example, text mining is a useful skill for faculty in the Social Sciences, the Humanities, and Language/Literature departments to learn, but at present, there is no program available to train all those interested or to help facilitate an interdisciplinary community of practice. Disciplinary prerequisites limit students' access to courses covering broadly applicable methodologies. Disciplinary boundaries also reduce opportunities for departments to collaborate on training for shared methodologies.

Lack of on-demand training: Students and faculty often do not know which methodology they need until the moment they need it. That moment is an opportunity for immersive, highly effective learning, but we currently have inadequate capacity and no curricular flexibility to

[1] Trevor Owens, "Defining Data for Humanists: Text, Artifact, Information, or Evidence?" *Journal of Digital Humanities* 1:1 (Winter 2011).

provide on-demand training.

Physical barriers: As technology hubs like Silicon Valley suggest, place matters. Staying current in the rapidly evolving world of data analysis and digital methods depends on face-to-face communications, and Middlebury needs more physical learning spaces that facilitate these interactions. Spaces equipped for teaching courses using data and digital methods are in short supply and currently do not allow for easy transitions between theory and practice. We need spaces where faculty can lecture on a concept and students can immediately use computers and software to explore its application. Teaching extends beyond the classroom, too. Faculty need well-designed offices where they can also teach, rather than crowding students around a single small computer monitor and then sending them out to find a lab computer where they can practice. Students need spaces most appropriate to their work and learning, especially in the evenings. Whether they are engaged in collaborative group projects, seeking advice and support from interns trained in digital tools, or getting help from "quantitative tutors" during drop-in hours, common areas with the right equipment, furniture, and flow are important. Shared spaces like these are a crucial feature of the most technologically innovative companies (and educational institutions) around the world.

Stagnant Funding for the DLA: In 2014, Middlebury received a three-year Mellon grant to launch the Digital Liberal Arts Initiative (DLA), a program investing in faculty development in digital scholarship. Over the past five years, more than 70 Middlebury College faculty across numerous disciplines have participated in DLA-funded programs supporting research projects and related efforts to expand the role of digital technology across the liberal arts. The DLA's impact also extended into pedagogy with a three year (2017-2020) grant from the Davis Family Foundation, "Building the Digital Capstone." Eight faculty members in four humanities and social science departments, with direct staff support from the DLA, developed a cluster of courses across a range of disciplines designed to prepare students to undertake senior work driven by digital questions, methods, and presentations.

After the Mellon grant expired in 2017, the DLA continued its operations year-to-year, with a new director each fall and an annual struggle to secure long-term funding and staffing. The 2017-2020 Davis Grant provided welcome and valuable support to pedagogical activities, but the future of the DLA has remained tenuous despite its effective and far-reaching support of faculty and students. It is time for Middlebury to develop a sustainable plan to continue to support and expand the DLA's goals and activities.

Declining enrollments in the Humanities: The Axinn Center for the Humanities at Middlebury released a report in July 2020 detailing the "state of the humanities" at Middlebury College. Co-directors Febe Armanios and Marion Wells describe declining enrollments and degrees in the humanities, diagnose the sources of these trends, and propose

new approaches to respond creatively and robustly while underscoring the centrality of the humanities to a liberal arts education. Among many creative and rigorous innovations, the Humanities Steering Committee writes that "we plan to work with Middlebury's Office of Digital Learning and Inquiry (DLINQ) and any new Digital entities currently in the planning stage [a.k.a. MiddData]…to promote novel interrogations of digital, computational, and data-oriented media within pedagogical and research pursuits" (p. 30). MiddData is excited to work with the Humanities Steering Committee to further illustrate and support core humanistic skills and dispel any notion students might have that a passion for the humanities conflicts with a "practical necessity" of acquiring the "hard skills" of data science. In particular, the report's call for exploration of interdisciplinary work and public humanities are promising directions for collaboration. The tools of data science and the digital humanities are useful to students in all disciplines as they pursue their passions, engage in independent inquiry, and solve problems.

The Proposal: What will MiddData do?

MiddData is a new initiative led by two faculty heads who will integrate and streamline existing efforts in digital arts, languages, and humanities with those in data science and qualitative and quantitative empirical research. The broad goal of MiddData is to support and expand data analysis and the use of digital tools and methods in teaching and research across Middlebury, and to do so in an efficient and equitable manner. The components of the proposal are detailed below. What follows is a long-term vision for MiddData. Given the current moment, these goals will be prioritized and implemented as resources allow.

1. House new campus-wide courses

a. Expand the current campus-wide Data Science course covering data acquisition, management, and visualization to include topics relevant to many disciplines such as web scraping, interactive web applications, text analysis, and machine learning. This expanded course, which will fulfill the DED requirement, will be accessible to first-year students in any field and will require no prior knowledge.

b. Introduce a campus-wide Introductory Statistics course covering the basics of statistical analyses, including hypothesis testing, ANOVA, and linear regression. This course will be staffed by multiple faculty across disciplines, be accessible to incoming first-year students with no prior experience, fulfill the DED requirement, and have the potential to prepare students for advanced courses in multiple disciplines. Departments would choose whether to count this course towards major requirements and whether to use it to replace Introductory Statistics courses currently offered within many departments.

c. Add new modular short courses and workshops in the fall and spring semesters that offer an introduction for students, faculty, and staff to the basics of software packages and/or programming (e.g., R, Gephi, Stata, ArcGIS, NVivo, Python, Omeka, Scalar), or to particular skills (e.g., data management practices, empirical writing, making choropleths, text encoding, photogrammetry, navigating the IRB, designing an RCT, filing a pre-analysis plan, podcasting). These courses can help meet the desire for on-demand training and may have prerequisites or corequisites.

d. Create two, credit-bearing boot camps for students who may benefit from additional support integrating data analysis into their coursework and career goals.

    i. A summer program targeting first-year students from populations underrepresented in data-intensive courses and majors would prepare them for success in introductory quantitative classes across the curriculum.

ii. An intensive Winter Term course, which we will pilot in 2021, that combines morning lectures introducing the general tools and practices of data science and digital methods with afternoon breakout sections where students apply their new skills to disciplinary projects.

2. Expand data and digitally oriented courses across the curriculum

a. Incentivize and facilitate the creation of "connector courses," at all levels and across the curriculum, that connect discipline-specific inquiry to the methodological tools taught in the new campus-wide courses supported by the *MiddData* program. Many such courses already exist, including those taught by members of the Ad Hoc Committee, by DLA leaders and fellows, and by Davis Grant Digital Capstone grant recipients.

b. Incentivize faculty to create first-year seminar connector courses that afford students immediate opportunities to get their feet wet with "real" data and digital methods from the moment they arrive on campus. These courses could have Data Science or Introductory Statistics as a corequisite, allowing interested first-year students to dive into tools and applications in small courses at the beginning of their academic careers.

c. Allow faculty and students to receive teaching or course credit for inquiry-based workshops. Students may, for instance, pursue a research project with a faculty member over a span of time, for varying credit hours. Faculty can use such courses to assemble teams of interdisciplinary experts.

3. Create and oversee new minors

a. Create a new Data Science Minor that combines courses in math/statistics, programming, and domain expertise. These courses will include Introduction to Data Science, foundation courses in mathematics, statistics and computing, and applied electives selected from a large menu of empirical courses offered across the curriculum.

b. Convene a working group to explore other possible minors distinct from Data Science. Possibilities include, in partnership with the Axinn Humanities Center, courses or an eventual minor in Public Humanities, Digital Arts and Humanities, and/or Medical Humanities. The Axinn Center for the Humanities is working on a proposal for a Public Humanities Initiative which has important ties with Digital Humanities. The Public Humanities Program aims to "help faculty and students produce public-facing and/or community-curated projects that address social and cultural issues of urgent

importance."[2] The Axinn Center directors look forward to working with Midd Data faculty and staff to develop this proposed program, and others. This provides another potential pathway to an academic minor or cluster.

4. Hire new faculty and staff

a. Appoint two faculty heads who will be responsible for the coordination, integration, and development of MiddData's programs and initiatives. The first will come from a background in statistics, computer science, or related field. The second will come from a digital arts and humanities background and will build on the work of the DLA.

b. Appoint or hire staff to work with the faculty co-directors in managing MiddData. Faculty co-directors, working with staff, will continue the successful communication and event-planning that the DLA has overseen in past years, including an up-to-date website, newsletters, blogs, workshops, guest speakers, and project presentations.

c. In academic departments and programs:

   i. Add new lines for faculty who can teach Data Science. These are likely to be housed in the Math and/or Computer Science Departments.

   ii. Incentivize departments to consider junior-level hires who can integrate data-centered and/or digitally-oriented courses with traditional curricular offerings. Such a contribution might be direct—by offering a course housed in MiddData—or indirect, by offering a connector course within their own discipline.

5. Leverage faculty resources across campus

a. Identify existing faculty, as well as new hires, who can contribute to the Introductory Data Science course, as well as possible courses in the Digital Arts and Humanities, modular courses, connector courses, and domain courses for the new Data Science and possible Digital Arts & Humanities minors/certificates.

b. Give faculty one-time teaching credit for taking Data Science or other appropriate courses. This will help more faculty across the curriculum become familiar with the program and develop new skills to innovate their teaching and research.

c. Make faculty salaries competitive, particularly for faculty in high demand due to their

[2] Febe Armanios and Marion Wells, "State of the Humanities Report at Middlebury College," Axinn Center for the Humanities at Middlebury (2020), p. 61.

methodological skills.

d. Expand the successful DLA faculty leave program into a MiddData and Digital Methods program to encourage faculty to explore both data-related and digital scholarship/pedagogy projects during their leaves.

6. Lower walls between disciplines

a. Identify courses across the curriculum that might count towards the Data Science and possible Digital Arts and Humanities or Public Humanities minors.

b. Encourage departments and programs to re-examine their empirical offerings and consider whether they can reduce prerequisites, cross-listed courses, or take other steps to make upper-level courses more accessible to students from outside their majors, including students pursuing a Data Science credential.

c. Incentivize faculty to design and cross-list new and interdisciplinary courses. For example, offer faculty extra teaching credit or a financial bonus if they cross-list with Midd Data programs or create a "connector course."

d. Host free weekly "lightning lunch" series at which faculty gather for short presentations of new research or pedagogy rooted in digital methods or data analysis.

e. Support departments interested in hiring faculty with expertise in data-driven and/or digital scholarship in disciplines with less experience in those approaches. This includes interdisciplinary mentorship, advising for start-up and technology costs, researching standards for tenure and promotion, funding team-teaching or faculty development, and providing leadership for potential "culture clashes" that might emerge.

7. Create and support professional development opportunities

a. Continue the core programs of the Digital Liberal Arts Initiative, including consultations, grants, and support for projects by faculty, staff, and students. This will include a continuation of a fellowship program in support of faculty on leave pursuing projects related to data and digital methods. Administration of the previous DLA fellowship program will shift to the MiddData Steering Committee.

b. Create a data and digital skills Winter term workshop and summer institute for faculty and staff, modeled on digital humanities summer institutes. These courses will prepare faculty and staff to integrate data analysis into teaching in their disciplines, which will in turn inspire students to take introductory-level courses in other data-driven

disciplines.

c. Offer funds for faculty who wish to travel to attend training workshops such as the University of Victoria Digital Humanities Summer Institute or ICPSR Summer Program, or to take online courses.

d. Expand the successful DLA Summer Institute program to include topics beyond Videographic Criticism, allowing Midd faculty/staff to instruct grad students and faculty from other institutions on specialized digital & data-driven methodologies in an immersive workshop model.

8. Increase support for the analysis of empirical data and digital methods

a. Expand financial support for both student and faculty projects to facilitate field research, experiments, and data acquisition.

b. Offer faculty financial or teaching credit incentives to employ student research assistants.

c. Expand and further support student "data and digital tutors" who offer methodological support to students and faculty engaged in research.

d. Expand and organize programs of student "data and digital consultants" who can assist faculty members in short and long-term research and curricular projects.

e. Expand and integrate staff from the Library, DLINQ, and ITS and associated resources for quantitative and qualitative data support and digital methods, including acquisition, access, instruction, and preservation.

f. Explore opportunities for inter-institutional collaboration that would allow us to share expertise and infrastructure with peer institutions.

9. Build new places for collaboration

a. Construct a new academic building as the hub of the MiddData initiative that will:

   i. House centrally-involved parties such as the Computer Science department as well as faculty teaching data and digital centered courses, while commiting to include a diverse group of digital and data scholars connected across existing disciplinary silos.

   ii. Provide better-equipped teaching spaces, including flexible computer classrooms, offices, and common areas with computational equipment and projection that may also be used by departments and faculty not be housed in

the new building.

iii. House staff, including data and digital librarians, computing support staff, and student quantitative tutors and DLINQ interns.

iv. Feature displays, exhibits, and “glass walls” to showcase collaborative work.

v. Provide computer lab spaces and common areas that draw students, faculty, and staff from across campus for engaged and collaborative work.

vi. Be home to new events, initiatives, and seminars in relevant areas, such as data science and the digital arts and humanities.

vii. Offer more public computers or loaner laptop computers that students can use to access licensed software in the same locations where their professors are located.

10. Improve computing infrastructure

a. Expand the current high-performance computing (HPC) cluster to lower barriers for students and faculty with high computational needs (analyzing large datasets, training deep networks, visualizing complex data, etc.). Experience from the current hiring season reveals that this cluster is an important recruiting tool that must be maintained.

   i. Grow the cluster by adding new "condo" nodes (e.g. funded with startup funds).

   ii. Budget for the maintenance and replacement of hardware.

   iii. Budget to cover electricity, storage & cooling expenses for the HPC cluster.

b. Expand and support cloud-based computing resources.

c. Budget for software and licenses for the HPC cluster and classroom/research computers.

d. Expand infrastructure for storing large amounts of data, both locally (via large disks) and in the cloud (via fast, high-bandwidth network connections).

Governance

The Provost's Office will appoint two faculty heads of MiddData, one with a background in the digital humanities and the other with a background in data science, statistics, or a related field. Working with advice from the Provost's Office, the faculty heads will form a steering committee of faculty and staff from diverse and relevant disciplinary backgrounds. In the first year of the initiative, this steering committee is charged with prioritizing and refining implementation of the proposal in collaboration and conversation with the broader community. The work of the larger steering committee and its connection to the rest of Middlebury will be guided by an executive committee composed of the two faculty directors and the dean of the library. This group will establish agendas for the steering committee, develop and monitor budgets, develop communications, and track progress. In addition, planning for the infrastructure required by MiddData will be integrated into larger Middlebury-wide planning efforts through the MiddLab team, which is comprised of representatives from ITS, DLINQ, CTLR, and the library.

The first academic year of the initiative, 2020-2021 takes place in the context of a global pandemic and reductions in institutional resources. The MiddData Steering Committee will reach out to colleagues across Middlebury to collect information and feedback and invite participation. The committee will identify and implement activities that balance impact with the constraints of the moment. So long as an institution-wide hiring freeze remains in effect, the committee will not seek to hire new faculty or staff for the initiative. However, the steering committee will use this time to identify medium and long-run staffing needs, establish medium and long-run plans and priorities, provide low-cost professional training and support to prepare faculty and staff to participate in the initiative, and support low-cost innovations in pedagogy and scholarship.

Conclusion

Empirical research and critical digital fluency have emerged as two of the strongest themes in Middlebury's Envisioning process. As Middlebury prepares for the decades ahead, the MiddData initiative will deepen and expand scholarship and learning across disciplines and the institution. By adopting these recommendations, Middlebury can not only change, restructure, and adapt, but also lead higher education as we continue to demonstrate that the liberal arts remain as relevant as ever in our rapidly-changing world.

Appendix A

Background and Process

In Spring 2018, the Provost's office issued a request for feedback and ideas related to Middlebury's strategic planning process. In response to this request, a number of departments and programs submitted documents highlighting the role that the collection and analysis of data plays in their work. Many expressed a desire to further support and develop empirical research and pedagogy at a moment when the role of data and data analysis is rapidly expanding in many of our disciplines, as well as in the world at large.

The College's Faculty Strategy Committee read and discussed these documents in Fall 2018, and then organized two open meetings in January 2019 to invite all members of the college community to further discuss initiatives related to the collection, visualization, and analysis of data involving methods not widely embraced in various disciplines right now. In addition to the four members of the Faculty Strategy Committee, these meetings were attended by 28 faculty members representing Biology, Computer Science, Economics, Food Studies, Geography, History, IP&E, Linguistics, Math, Physics, Political Science, and Sociology, and by 17 staff representing academic affairs, CTLR, DLA, DLINQ, Experiential Learning Centers, ITS, the Library, the Provost's office, and the Registrar's office. Attendees at these meetings stressed the importance of data literacy to our students and staff, highlighted great work that is already being done at Middlebury, identified key challenges and opportunities, and expressed enthusiasm for a new initiative that could support and expand these activities.

Following these meetings, in February 2019 the Faculty Strategy Committee appointed the members of the first Ad Hoc Committee on Empirical Teaching and Research: Bill Koulopoulos, Director of Technology and Learning Spaces; Sarah Laursen, Assistant Professor of History of Art and Architecture and Curator of Asian Art; Matthew Lawrence, Assistant Professor of Sociology and Director of Privilege and Poverty; Alex Lyford, Assistant Professor of Mathematics; Caitlin Myers, Associate Professor of Economics and a representative of the Faculty Strategy Committee; Daniel Scharstein, Charles A. Dana Professor and Chair of Computer Science; Wendy Shook, Science Data Librarian; and, Mark Spritzer, Associate Professor and Chair of Biology.

The Ad Hoc Committee's charges were: (1) evaluate current challenges and opportunities in the areas of digital humanities, data science, and qualitative and quantitative empirical research, and (2) develop a proposal for a new initiative to expand access to sophisticated data-based methodologies.

The Ad Hoc Committee met weekly between February and May 2019 to develop

recommendations. The Committee reviewed and discussed materials generated at the open meetings, as well as practices and programs at other colleges and universities. In addition, members of the Committee met and consulted with members of the following departments and programs: Biology, Computer Science, Economics, Geography, Film and Media Studies, History of Art and Architecture, Linguistics, Mathematics, Molecular Biology and Biochemistry, Neuroscience, Political Science, Physics, Psychology, and Sociology and Anthropology. Members of the Committee also discussed the proposal with the President of MIIS, Dean for Faculty Development and Research, and with representatives of CTLR, DLA, DLINQ, the Library, and ELC. The Committee presented at the April Plenary Session of the Middlebury Faculty, to the President's Leadership Council, and to the Board of Trustees Strategy Committee. In May of 2019 the Ad Hoc Committee submitted a detailed proposal for a MiddData initiative to the Faculty Strategy Committee and Provost's office.

After reviewing this proposal, the Provost's Office and Faculty Strategy Committee responded enthusiastically, but also noted the need to better integrate the MiddData initiative with the DLA. In Fall 2019 the Provost's office brought these two streams together, forming a second ad hoc committee to more fully integrate the work that the DLA began in 2014 with the proposal developed as part of the Envisioning Process. Members of this second group included: Carrie Anderson, Assistant Professor of History of Art and Architecture; Amy Collier, Associate Provost for Digital Learning; Sarah Laursen, Assistant Professor of History of Art and Architecture; Matt Lawrence, Assistant Professor of Sociology; Alex Lyford, Assistant Professor of Mathematics; Vijay Menta, Assistant Vice President and Chief Information Officer; Jason Mittell, Professor and Chair of Film and Media Culture; Amy Morsman, Professor of History and Director of the DLA; Caitlin Myers, Professor of Economics and Member Faculty Strategy Committee; Phil Murphy, Associate Professor at MIIS and Director of the META Lab; James Ralph, Dean of Faculty Development and Research and Rhenquist Professor of American History and Culture; Mike Roy, Dean of the Library; and Daniel Scharstein, Charles A. Dana Professor of Computer Science and Chair of Computer Science. That second group spent time revising the proposal, considering questions about how to move forward with the proposed activities, how to communicate with and gather new ideas from the rest of the community, as well as questions having to do with infrastructure.

Following this process, in the summer of 2020 three colleagues continued to refine and refine the proposal: Kathryn Morse, John C. Elder Professor of History and Environmental Studies and new faculty director of the DLA; Caitlin Myers, John G. McCullough Professor of Economics, member of the Faculty Strategy Committee, and member of both ad hoc planning committees; and Mike Roy, Dean of the Library and member of the second ad hoc planning committee. Kathryn, Caitlin, and Mike refined and finalized the proposal in consultation with members of both ad hoc planning committees, the DLA, and the Humanities Steering

Committee. This document is the ultimate product of this lengthy process.

Naming the initiative

Throughout the planning process, both the first and second ad hoc planning committees devoted substantial time and energy to discussing the name of the new initiative. The committees quickly came to understand that the language we use to describe the tools and methodologies the initiative targets varies by discipline, with colleagues in the humanities, languages, and arts often referring to "Digital Humanities" or "Digital Liberal Arts," colleagues in computer science and mathematics to "Data Science," and colleagues in the natural and social sciences to "empirical methods," "econometrics," "experiments," and "statistics." While we make no claim that these terms are synonymous, it became clear in the discussion that the overlap between these methodological descriptors is large and that there are meaningful pedagogical and scholarly synergies and efficiencies to be found in their intersections.

In choosing a name for the new initiative, the ad hoc committees wished to simultaneously signal the breadth and inclusivity of the initiative to faculty communities, signal its promise and potential to constituencies who may be most familiar with the popular term "data science," and choose a name that is simple and punchy. The committees discussed and debated and brainstormed names through many meetings, and ultimately decided to pair a longer and more inclusive name for the initiative—The Middlebury Initiative for Data and Digital Methods—with a shorter nickname—MiddData. The committees hope that the longer name signals the breadth of the initiative while the nickname rolls off the tongue and will over time and through the work that we do, come to connote the central theme of data and digital fluency.